Renewable and Sustainable Energy Reviews 28 (2013) 749–756

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Decomposing global solar radiation into its direct and diffuse components

John Boland *, Jing Huang, Barbara Ridley

School of Mathematics and Statistics, Barbara Hardy Institute, University of South Australia, Mawson Lakes Boulevard, Mawson Lakes, SA 5095, Australia

ARTICLE INFO

Article history:
Received 3 December 2012
Received in revised form
5 August 2013
Accepted 11 August 2013
Available online 6 September 2013

Keywords:
BRL model
Direct normal radiation
Diffuse radiation
Statistical modelling

ABSTRACT

To assess the viability of proposed solar installations, knowledge of global solar radiation is not sufficient. For stationary photovoltaic plant, we require global radiation series, but also the contemporaneous diffuse radiation series. Alternatively, for concentrated solar thermal, we need global and direct normal solar radiation. In this paper, we investigate whether one can simply use a model for predicting diffuse radiation using multiple predictions derived by our research team, the Boland–Ridley–Lauret (BRL) model, to give delineations of both diffuse and direct or if we need to use another model for direct or develop a new direct normal statistical model.

Contents

1. Introduction

The evaluation of the performance of a solar collector such as a solar hot water heater or photovoltaic cell requires knowledge of the amount of solar radiation incident upon it. Solar radiation measurements are typically only for global radiation on a horizontal surface. They may be on various time scales, by minute, hour or day. Additionally, one can infer daily totals from satellite images. These global values comprise two components, the direct and the diffuse. DNI, “the direct normal irradiance, is the energy of the direct solar beam falling on a unit area perpendicular to the beam at the Earth's surface. To obtain the global irradiance the additional irradiance reflected from the clouds and the clear sky must be included” [1]. This additional irradiance is the diffuse component.

For various applications, one needs knowledge of diffuse solar radiation and for others, one needs to have measured or estimated values of direct solar radiation. For flat plate collectors and house energy analysis, we require global and diffuse radiation series but for concentrated solar thermal, we need global and direct solar radiation. If only global radiation on a horizontal surface is available through measured data or inferred from satellite images, one will need some type of model to estimate either the diffuse or direct from the global values. When research first began on this topic, the solar collectors in use were all flat plate, and so attention was focused on developing diffuse radiation models.

There is an added reason for computing values of the diffuse radiation. Typically solar collectors are not mounted on a horizontal surface but tilted at some angle to it. Thus it is necessary to calculate values of total solar radiation on a tilted surface given

* Corresponding author. Tel.: +61 8 83025781.
E-mail address: john.boland@unisa.edu.au (J. Boland).

1364-0321/$ - see front matter
http://dx.doi.org/10.1016/j.rser.2013.08.023

values for a horizontal surface. It is not possible to merely employ trigonometric relationships to calculate the solar radiation on a tilted collector. This is because the diffuse radiation is anisotropic over the sky dome and the "radiative configuration factor from the sky to the tilted solar collector is not only a function of the collector orientation, but is also sensitive to the assumed distribution of the diffuse solar radiation across the sky" [2]. There are two different approaches to calculating the diffuse radiation on a tilted surface; using analytic models, for example the Brunger approach [2] or empirical models such as the BRL model [3]. Each rely on knowledge of the diffuse radiation on a horizontal surface. The diffuse component is not generally measured. Consequently, a method must be derived to estimate the diffuse radiation on a horizontal surface based on the measured global radiation on that surface.

Numerous researchers have studied this problem and have been successful to varying degrees. Liu and Jordan [4] developed a relationship between daily diffuse and global radiation which has also been used to predict hourly diffuse values. The predictor typically used in studies is not precisely the global radiation but the "hourly clearness index k_t, the ratio of hourly global horizontal radiation to hourly extraterrestrial radiation" [5]. Orgill and Hollands [6] and Erbs et al. [7] correlate the hourly diffuse radiation with k_t, but Iqbal [8] extended the work of Bugler [9] to develop a model with two predictors, k_t and the solar altitude. Reindl et al. [5] use stepwise regression to "reduce a set of 28 potential predictor variables down to four significant predictors: the clearness index, solar altitude, ambient temperature and relative humidity." They further reduced the model to two predictor variables, k_t and the solar altitude, because the other two variables are not always readily available. Another possible reason was that some combinations of predictors may produce unreasonable values of the diffuse fraction, eg. greater than 1.0 [5]. Skartveit et al. [10] developed a model which in addition to using clearness index and solar altitude as predictors, also added a variability index. This is meant to add the influence of scattered clouds on the sky dome. As well, Gonzales and Calbo [11] stress the importance of including the altitude and the variability of the clearness index in any predictions of the diffuse fraction. Aguiar [12] fitted an exponential model to Mediterranean daily data using only the clearness index and found a consistency of fit amongst locations of similar climate.

Boland et al. [13] presented the use of a decaying logistic function to estimate the diffuse fraction from knowledge of the clearness index. Subsequently, the lead author of that paper combined with other researchers to provide a theoretical basis for selecting that form of the model [14]. This concept was further developed by adding more predictor variables to enhance the fit, resulting in the Boland–Ridley–Lauret (BRL) model [3]. The modelling effort in these three studies can be classified as from a frequentist approach to statistical modelling. This refers to the classical least squares estimation procedure that was used to perform the parameter estimation. In related work [15,16], the problem was undertaken using an alternative statistical starting proposition, Bayesian model building and parameter estimation. It was reassuring that using two separate modelling approaches, the same predictor variables were found to be significant and the parameter estimates proved to be very similar.

In recent years, there has been increasing interest in both concentrating solar thermal (CSP) and concentrating solar photovoltaic (CPV) installations, and as a consequence, an increasing interest in reliable estimation of direct normal radiation. So, we now have the situation where for some applications, we need to estimate diffuse radiation from global radiation, and for others, direct normal radiation (DNI) from global radiation. As testimony to this, Perez–Higueras et al. [17] have developed a simplified model to predict direct normal from global. Additionally, the latest version of Meteonorm software [18] includes two models in this area, one statistically based model, the BRL model [3] for estimating diffuse from global, and one physically based model, the Perez model [19], to estimate DNI from global.

The question that comes immediately to mind is whether we need a plethora of models, specifically do we need a "best" model for estimating diffuse from global and a "best" model for estimating DNI from global? Or, can one model suffice, wherein estimation of the diffuse from global is performed, for instance, and then the DNI is calculated from the other two components? In this paper, we will provide evidence that using the BRL model [3] to estimate diffuse solar fraction, and from it calculate DNI performs as well as any present model specifically designed to estimate the DNI from knowledge of the global. The implication is that we do not need another complex model to model direct solar radiation, because the direct solar radiation coming from the modelling of diffuse solar radiation is sufficient.

The paper is organized as follows. Section 2 describes the development of the logistic function model of hourly direct normal solar radiation with multiple predictors. Comparison of the logistic function model with other models and error analysis is given in Section 3. How direct normal solar radiation is calculated from the BRL model for modelling diffuse solar fraction with multiple predictors and comparison of this procedure with other models is described in Section 4. The final section is devoted to conclusions.

2. Historical development of the diffuse fraction model

The original approach to diffuse fraction estimation from the clearness index relied on a basic assumption that there are three separate regions in the scatterplot – Fig. 1, reflecting differing processes. Lanini [20] discusses this with using the Reindl model [5] to as an example. The model is given below:

$$d=\eta_1+\gamma_1 k_t+\delta_1 \sin\alpha \quad 0\le k_t\le 0.3 \quad d\le 1.0$$
$$d=\eta_2+\gamma_2 k_t+\delta_2 \sin\alpha \quad 0.3<k_t<0.78 \quad 0.1\le d\le 0.97$$
$$d=\eta_3+\gamma_3 k_t+\delta_3 \sin\alpha \quad k_t\ge 0.78 \quad 0.1\le d$$

Lanini shows that for an example data set, the diffuse fraction varies in the middle sub-interval of $0.3\le k_t\le 0.78$, with solar altitude α but has very little variation in the end ranges $k_t<0.3$ and $k_t>0.78$. This is then used as the justification for breaking the interval for k_t into three segments and using separate models in each sub-interval. Reindl may well have been guided by earlier work where a similar splitting was done. Many of the earlier approaches, such as that of Orgill and Hollands [6], used piecewise

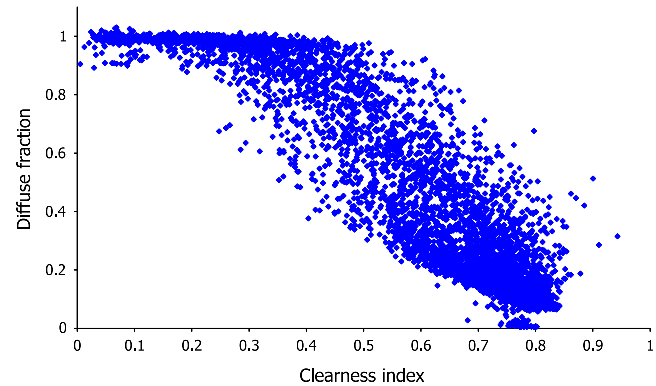

Fig. 1. Diffuse fraction versus clearness index for Adelaide.

linear models for the sections. There are, in our opinion, a number of problems with this splitting.

- Various authors use different end points for the sub-intervals. Orgill and Hollands [6] use 0.35, 0.75 while Erbs [7] uses 0.22, 0.8.
- The argument used above assumes that there are only two significant predictors. Reindl uses only two as they are the ones from his set of 28 that are most easily obtained for any location. This covering of the spread of data in the middle sub-interval by using the solar altitude works to an extent, but as we will see later, the use of other equally available predictor variables covers more of that spread and also the spread in the two sub-intervals at the ends.
- For the Reindl model per se, there are discontinuities at the boundaries of the sub-intervals.
- It must be stated that the third reason was not necessarily apparent to any of the early modellers. The adoption of a single model formulation for the whole interval for k_t enables the alteration of the model to suit climate change projections, much more easily than have fixed sub-intervals.

We now encapsulate the steps in the development of the BRL model. When we first began looking at the problem in preparation for the first version of the model [13], it seemed from a purely curve fitting perspective, that a type of decay function should be appropriate. This led to the idea to first construct the variation of the diffuse fraction as a moving average through the k_t interval. Such an exercise is depicted in Fig. 2. From this, it was relatively straight forward to select a decaying logistic function as the appropriate one. The next iteration was done in a much more systematic manner [14]. It was decided to transform the data to a form that is amenable to standard linear regression techniques. Note that for linear regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (1)$$

the assumption is that the x_i are known, and the y_i are random variables that are independent and identically distributed (iid). This means that the transformation should be of a type to result in a homogeneous band of variation in the dependent variable as the independent variable increases, as happens after the transformation – see Fig. 3. The data was now in a form suitable for modelling with a line of best fit – see Fig. 4, whereupon the data and line were back transformed to give the fit to the original data – see Fig. 5. This seemed a more mathematical approach to the problem than a simple moving average, and from it we felt we were justified in selecting a decaying logistic function to use. The final step in the development for the single predictor model was to perform the activity above for several locations. The parameter estimates were sufficiently similar to set in train the idea of combining data sets from the various locations and constructing a model that may be used for any location – see for example the model applied to another location in Fig. 6.

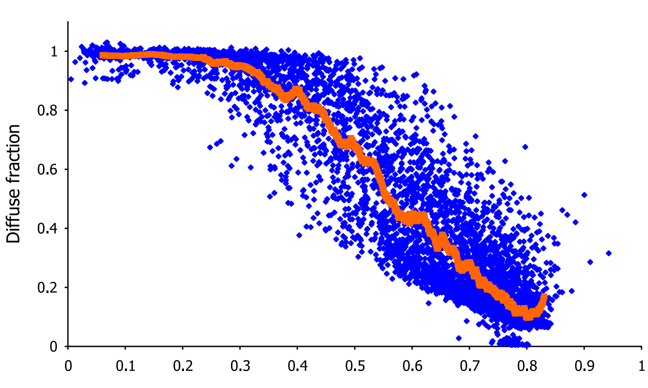

Fig. 2. Diffuse fraction versus clearness index for Adelaide with moving average superimposed.

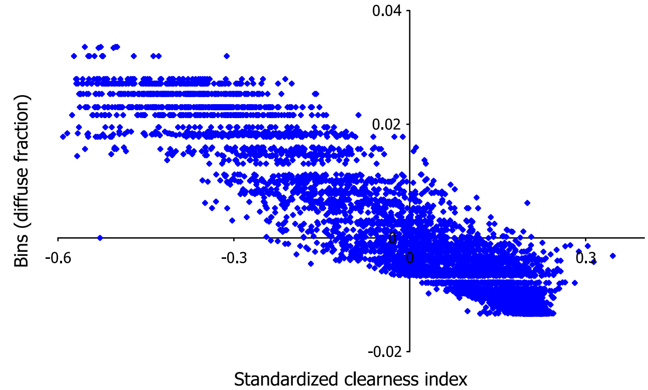

Fig. 3. Diffuse fraction versus clearness index for Adelaide transformed.

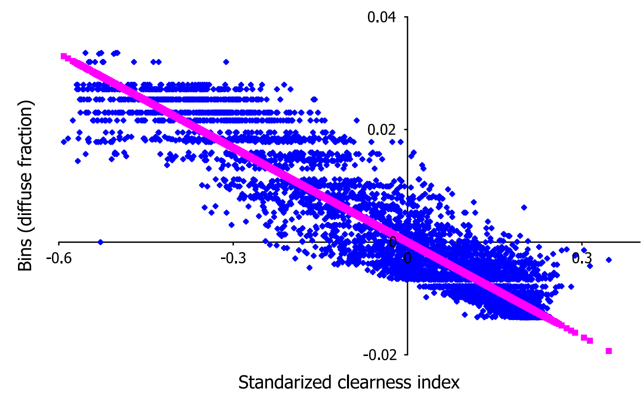

Fig. 4. Diffuse fraction versus clearness index for Adelaide transformed with line of best fit.

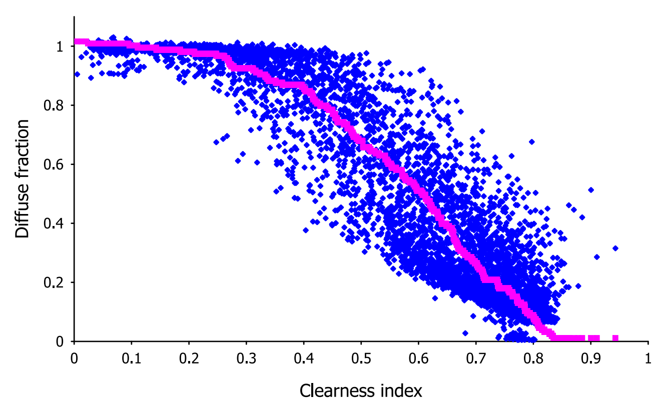

Fig. 5. Diffuse fraction versus clearness index for Adelaide back transformed with fit.

The final step in the process resulted in the BRL model [3,15,16]. What was added was four other predictor variables to cover much more of the spread of the data. These are apparent solar time *AST*, solar altitude angle α, daily clearness index K_t and persistence ψ. The first three are self-explanatory in terms of what they are. The solar altitude had been employed in various other models. The inclusion of AST reflects the fact that the atmosphere is generally more turbid in afternoon than morning. See [3] for more details. The result of adding these extra predictors is shown in Fig. 7, fitting the spread of the data much better.

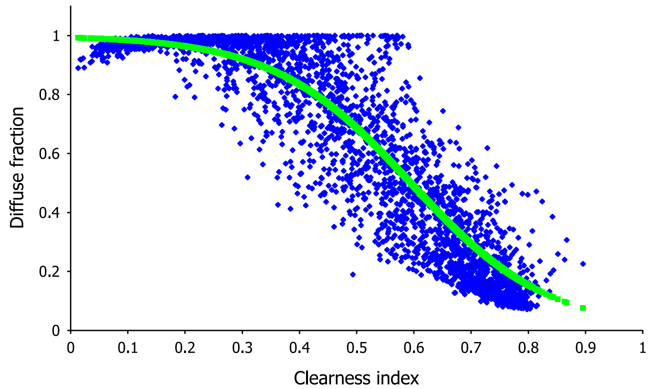

Fig. 6. Diffuse fraction versus clearness index for Geelong with fit.

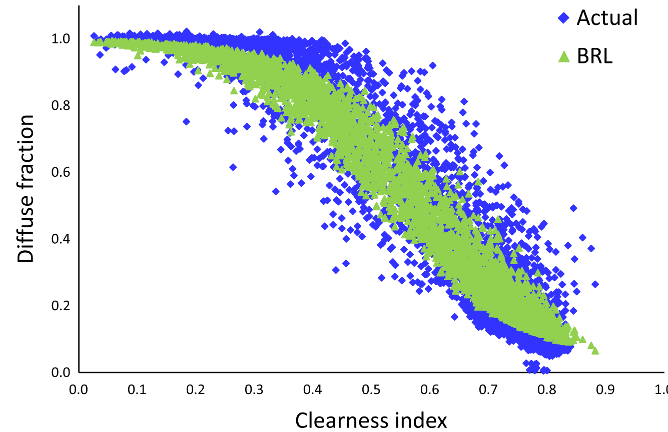

Fig. 7. The BRL diffuse fraction model fit to Adelaide data.

3. Logistic model

Boland et al. [13,14] found that using a decaying logistic function is a good way to model diffuse solar radiation and also it is widely used in ecology and species growth representations [21,22]. To use a logistic function for modelling direct normal solar radiation, it is growth with respect to clearness index.

According to Banks [23] and Jeffrey [24] using standard integration techniques, Thornley et al. [25] obtains a modified logistic function as

$$G_t = \frac{N \times M}{N + (M - N) \times e^{-r_0 \cdot t}} \tag{2}$$

Here, G_t represents population number, and N and M have biological meaning for populations with a strong interaction among individuals that controls their reproduction. If $N < M$, this represents logistic growth which is the situation needed here and if $N > M$, there is logistic decay. r_0 is the maximum possible rate of population growth. If r_0 is low this means a slow rate of growth, otherwise it will be fast. When applying this structure to modelling direct solar radiation, G_t is replaced by the direct normal solar radiation I_{DN}, r_0 is replaced by β_1 and t is replaced by the clearness index k_t. Ridley et al. [3] also suggest using four other important parameters for modelling the diffuse fraction which are apparent solar time AST, solar altitude angle α, daily clearness index K_t and persistence ψ. These are adopted here as well. The multiple predictor logistic model is

$$I_{DN} = \frac{N \times M}{N + (M - N) \times e^{-\beta_1 \cdot k_t - \beta_2 \cdot AST - \beta_3 \cdot \alpha - \beta_4 \cdot K_t - \beta_5 \cdot \psi}} \tag{3}$$

The data chosen to build the model is multiple location data [3] which is aggregated data from seven locations worldwide (Adelaide, Darwin, Bracknell, Lisbon, Macau, Maputo and Uccle from year 2001 to 2005, there are 7338 hourly diffuse radiation data points). The method of ordinary least squares in Solver (an optimization tool in EXCEL) is used to obtain all the parameter estimates, Eq. (3) becomes

$$I_{DN} = \frac{0.006 \times 4.38}{0.006 + (4.38 - 0.006) \times e^{-7.75 k_t - 1.185 AST - 1.05 \alpha - 0.004 K_t + 0.003 \psi}} \tag{4}$$

Eq. (4) is applied to four individual locations; Adelaide (from year 2003 to 2004, 4741 hourly diffuse radiation data points), Darwin (from year 2001 to 2005, 2597 hourly diffuse radiation data points), Lisbon (for year 1980, 3422 hourly diffuse radiation data points) and Mt Gambier (from year 1973 to 1977, 14,058 hourly diffuse radiation data points) to test the efficacy of the modelling of hourly direct normal solar radiation. The data obtained from the Bureau of Meteorology, Australia. Fig. 8 shows

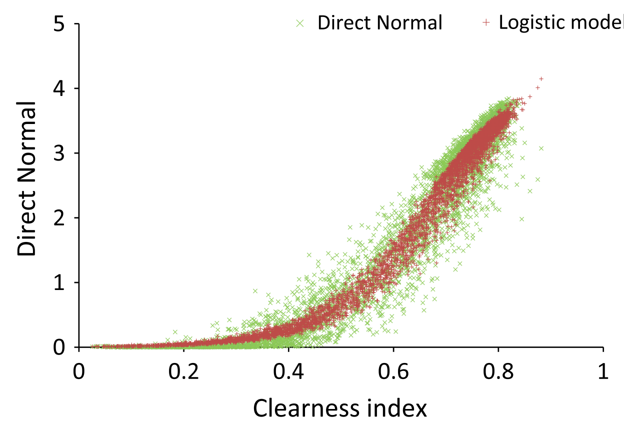

Fig. 8. The logistic model fit for direct normal data in Adelaide.

that the logistic model gives a good fit for the direct normal data for the southern hemisphere location of Adelaide. In the northern hemisphere location of Lisbon there is also a good fit as shown in Fig. 9.

4. Comparison with other models

There are many models used to predict direct normal solar radiation, but one of the most recognized is the Perez model. So, we will use it to compare with the logistic model.

The Perez model [19] is a four dimensional coefficient matrix model based on the Maxwell's model [26] for estimating direct normal solar radiation. The basic idea of the Perez model is to use a coefficient function $X(K_t^{'}, Z, W, \Delta K_t^{'})$ to improve the estimate values I_{disc} from Maxwell's model, as given in

$$I_{DN} = I_{disc} \cdot X(K_t^{'}, Z, W, \Delta K_t^{'}) \tag{5}$$

Here, $K_t^{'}$ is a zenith angle dependent expression of the clearness, Z is solar zenith angle, W is atmospheric perceptible water and $\Delta K_t^{'}$ is the stability index.

Fig. 10 shows the Perez model against the actual data in Adelaide and it is not performing well for higher clearness index values. The Perez model seems to have limitations for predicting the higher values of direct normal. The same problem also appears in the other three locations. Thus feature is a common problem for use of models for either diffuse or direct radiation that have been developed for Northern Hemisphere locations, when they are applied to Southern Hemisphere sites. It does also seem to be an issue for Lisbon – see Fig. 11.

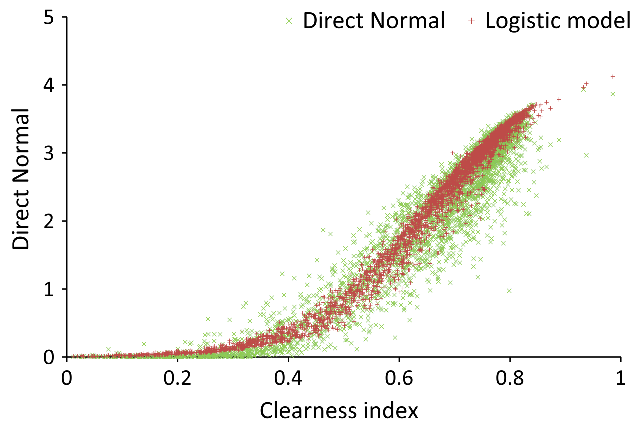

Fig. 9. The logistic model fit for direct normal data in Lisbon.

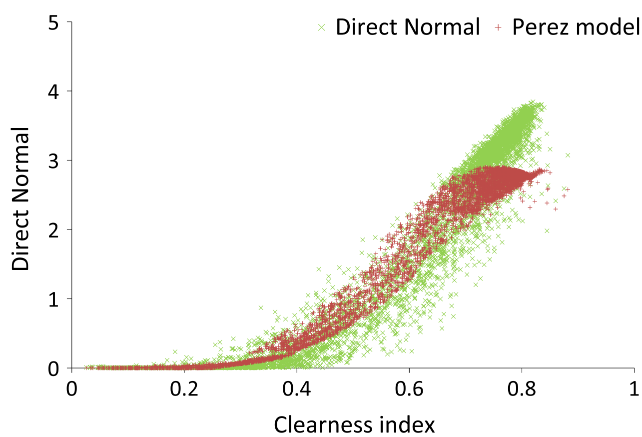

Fig. 10. The Perez model fit for direct normal data in Adelaide.

For the purpose of formal error analysis of the proposed models, the following measures are considered: median absolute percentage error (MeAPE), mean bias error (MBE), normalised root mean square error (NRMSE) and Kolmogorov–Smirnov test integral (KSI). MeAPE captures the size of the errors, while MBE is used to determine whether any particular model is more biased than another. NRMSE is a measure of overall model quality related to regression fit. What this means is that is how far the data deviates from the model. What is more informative is in essence how far the regression line is from the line $Y=X$, where the y's are the predicted values from the model, and x's are the data values. Interestingly, Willmott and Matsuura [27] produce convincing arguments as to why the mean absolute error (MAE) is a superior error measure to the RMSE. They argue that the RMSE is a function of three characteristics of a set of errors.

It varies with the variability within the distribution of error magnitudes and with the square root of the number of errors ($n^{1/2}$), *as well as with the average-error magnitude* (*MAE*).

KSI is a new model validation measure based on the Kolmogorov–Smirnov test [28] which has the advantage of being nonparametric. The KSI measure was proposed by Espinar et al. [29] to assess the similarity of the cumulative distribution functions (CDFs) of actual and modelled data over the whole range of observed values.

Definitions of all the measures are as follows:

$$MeAPE = MEDIAN\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\right)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$$

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}}{\bar{y}}$$

where $\hat{y}_i$ are predicted values, y_i are measured values and $\bar{y}$ are average of measured values.

$$KSI(\%) = 100 \times \frac{\int_{x_{min}}^{x_{max}} D_n \, dx}{\alpha_{critical}}$$

where x_{max} and x_{min} are the extreme values of the independent variable, and $\alpha_{critical}$ is calculated as $\alpha_{critical} = V_c \times (x_{max} - x_{min})$. The critical value V_c depends on population size N and is calculated for a 99% level of confidence as $V_c = 1.63/\sqrt{N}, N \geq 35$. D_n are the differences between the cumulative distribution functions (CDFs) for each interval. The higher the KSI value, the worse the fit of model to data.

Table 1 shows that the logistic model is better than the Perez model in all four error analyses at all locations, except MBE in Lisbon. This is further illustrated in Figs. 12 and 13 for the KSI which show observed and modelled CDFs, as well as differences between them over the whole range of the data. Clearly, the logistic model obtains estimates closer to the measured values and lower values of D_n. Thus, the logistic model appears at least as accurate as the Perez model for predicting hourly direct normal solar radiation. Therefore, there appears to be no advantage of using arguably the best performing Direct Normal model in the literature over using the multiple predictor direct normal model developed here.

Table 1
Results of error analysis of two models in four locations.

Error measure	Adelaide	Darwin	Lisbon	Mt Gambier
Logistic model				
MeAPE	9.20%	8.36%	10.92%	25.47%
MBE	−0.030	0.040	−0.122	0.094
NRMSE	14.20%	13.35%	16.27%	28.44%
KSI	22.27%	13.01%	26.11%	20.36%
Perez model				
MeAPE	20.94%	13.94%	18.18%	27.60%
MBE	−0.145	−0.231	−0.087	−0.139
NRMSE	24.81%	17.73%	22.35%	33.74%
KSI	80.04%	60.43%	47.70%	30.82%

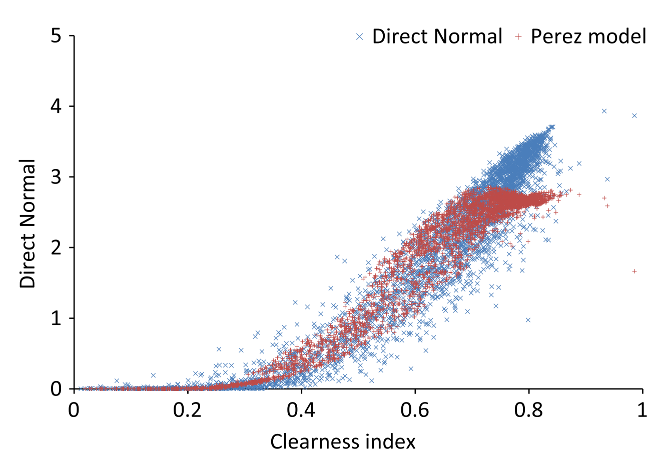

Fig. 11. The Perez model fit for direct normal data in Lisbon.

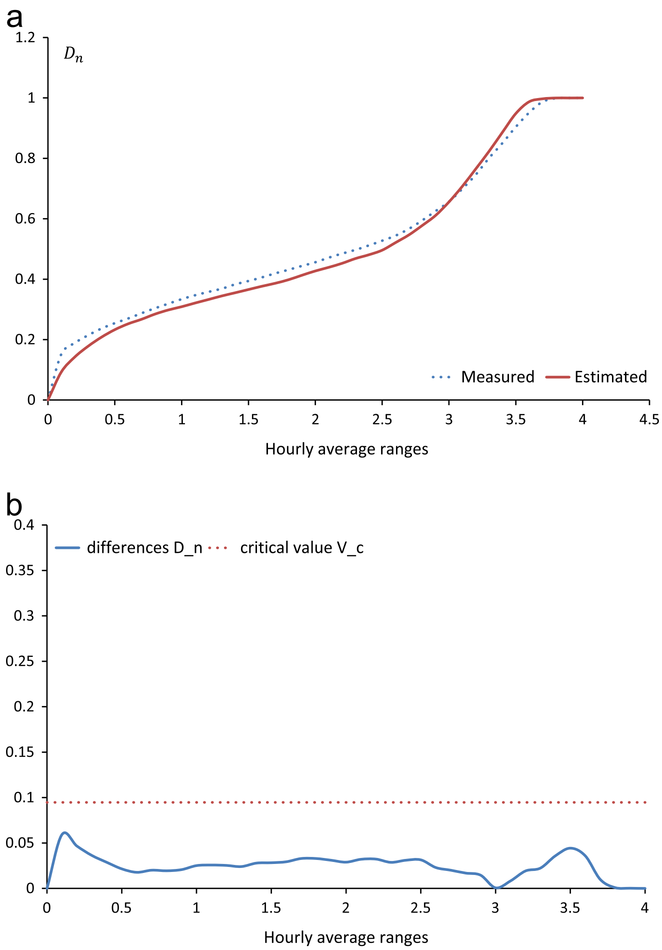

Fig. 12. Plot of the logistic model of CDF for the measured and predicted data sets (left) and the differences D_n between those (right) at Adelaide. The dotted line marks the critical value V_c.

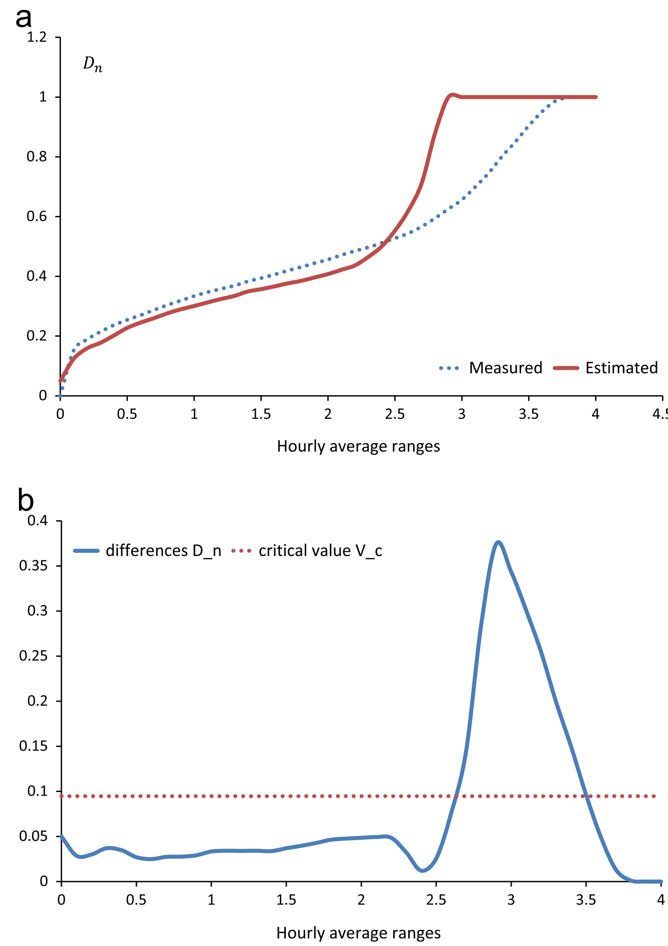

Fig. 13. Plot of the Perez model of CDF for the measured and predicted data sets (left) and the differences D_n between those (right) at Adelaide. The dotted line marks the critical value V_c.

5. Boland–Ridley–Lauret (BRL) model

The performance of the direct normal radiation model derived in Section 2 is compared with the direct normal radiation estimated through using the BRL model [3]. The purpose of this comparison is to ascertain whether using a model already reported in the literature, the BRL diffuse fraction model [3], is sufficient for estimating DNI from global through the chain global-diffuse-DNI. In Section 3, it was shown that the model in Eq. (4) performs at least as good as the best performing model in the literature. If the BRL model performs just as well, then there is no need for this new approach.

To obtain the direct normal solar radiation from the BRL model, the following steps will be applied.

First, use the BRL model given by Ridley et al. [3]

$$d = \frac{1}{1+e^{-5.38+6.63k_t+0.006AST-0.007\alpha+1.75K_t+1.31\psi}} \tag{6}$$

to obtain the diffuse fraction, d.

Second, using the following equation, we can calculate the direct normal solar radiation.

$$I_{DN} = \frac{I_G-(d\cdot I_G)}{\sin(\alpha)} \tag{7}$$

Here, I_G is global solar radiation, α is solar altitude angle and I_{DN} is direct normal solar radiation. Using Eq. (7), we can obtain the BRL model fit to the direct normal data in Adelaide which is shown in Fig. 13.

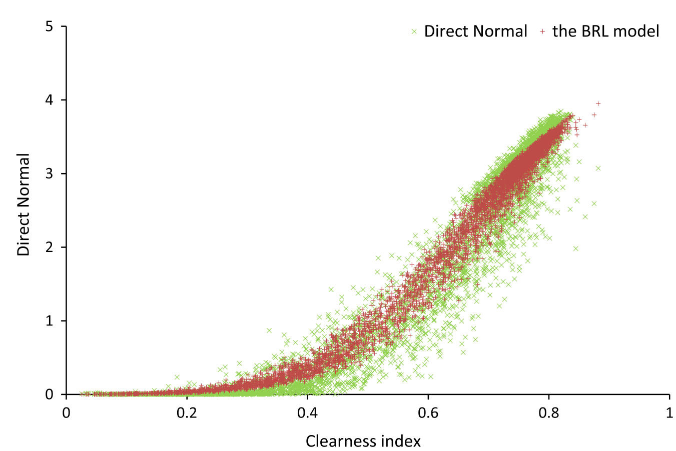

Fig. 14. The BRL model fit for direct normal data in Adelaide.

Comparing Figs. 8 and 14 shows that the BRL model results seem to cover the data more than the logistic model, but it is hard to see the difference between these two figures. So the same error analysis as before has also been used for the predicted values of the BRL model.

Table 2 shows that for all error measures the BRL model performs as well as the newly derived model and thus at least as well as or better than the Perez model. Since it performs slightly better than the logistic model in MeAPE and NRMSE we could say that the BRL model, in a 'local' sense, is better than the logistic

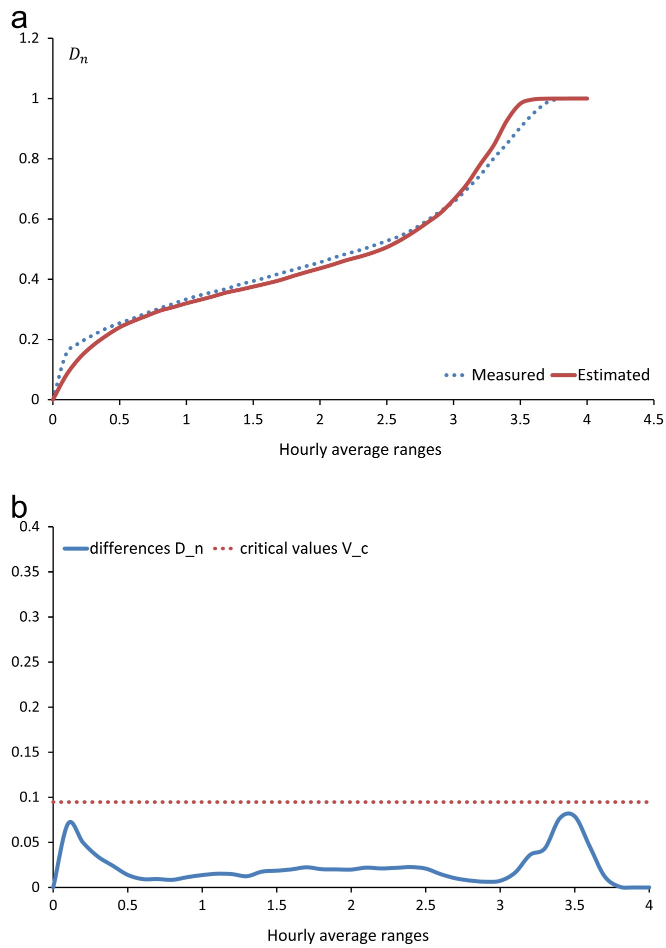

Fig. 15. Plot of the BRL model of CDF for the measured and predicted data sets (left) and the differences D_n between those (right) at Adelaide. The dotted line marks the critical value V_c.

Table 2
Results of error analysis of the BRL model in four locations.

Error measure	Adelaide	Darwin	Lisbon	Mt Gambier
Logistic model				
MeAPE	9.20%	8.36%	10.92%	25.47%
MBE	−0.030	0.040	−0.122	0.094
NRMSE	14.20%	13.35%	16.27%	28.44%
KSI	22.27%	13.01%	26.11%	20.36%
BRL model				
MeAPE	8.87%	8.12%	9.65%	24.90%
MBE	−0.068	0.049	−0.139	0.045
NRMSE	14.49%	12.71%	16.16%	27.98%
KSI	24.72%	13.39%	29.21%	10.64%

model. However, for error measures MBE and KSI, the logistic model is slightly better than the BRL model, except in Mt Gambier. It is an illustration that, in a 'global' sense, the logistic model developed here has better predictive ability, but not for all locations. Figs. 12 and 15 also show that instead of slightly different 'global' and 'local' performance, the BRL model obtains better predicted values for most of the ranges (in the middle ranges from 0.5 to 3) and the logistic model is a little bit better at the ends of the range (from 0 to 0.5 and 3 to 4). Therefore, based on their 'global', 'local' performance, D_n and the value of difference of D_n, it is concluded that the BRL model is a slightly better than the logistic model.

6. Discussion

The development of models for estimating diffuse solar radiation began in 1960. Liu and Jordan [4] used the diffuse solar radiation depending on different degrees of cloudiness or ranges of clearness index for 98 locations across United States and Canada. Numerous models have been presented for estimating diffuse radiation since then, such as Orgill and Hollands [6] who used a correlation function to estimate hourly diffuse radiation on a horizontal surface and Erbs et al. [7] who used a curvilinear function to establish a relationship between the hourly diffuse fraction and the hourly clearness index k_t. For the purpose of reducing the standard error of the correlation function, Reindl et al. [5] proposed 28 potential predictor variables for estimating diffuse fraction data and utilized stepwise regression reducing the 28–4 significant predictors: the clearness index, solar altitude, ambient temperature and relative humidity. In 1992, Perez et al. [19] developed a direct normal radiation model which is a four dimensional coefficient matrix model based on the Maxwell's model [26]. Until now, the Perez model has been recognized as one of the most accurate models for estimating direct normal radiation. In 2001, Boland et al. [13] developed a logistic function to estimate diffuse fraction which is unlike previous methods that used either piecewise linear or simple nonlinear functions. To further validate the logistic model, Boland et al. [14] outlined the theoretical development of the logistic function, for estimating diffuse solar radiation. Then, Ridley et al. [3] developed the Boland–Ridley–Lauret (BRL) model to improve the accuracy of the logistic function by adding more predictors which was then verified further by using a Bayesian approach to arrive at the same model structure [15,16].

Recently, in the literature, most papers about diffuse fraction study either test many models [30,38] or add their own correlations [32,33]. For example, Kudish and Evseev [31] evaluated four different correction models: Drummond [34], LeBaron et al. [35], Battles et al. [36] and Muneer and Zhang [37]. Through error analysis and scoring systems, such as the coefficient of determination of R^2, RMSE, MBE, Percentage average deviation (PAD), deviation (SD), t-statistic, accuracy score (AS) and Kudish and Rahima (K&R) [38], Kudish and Evseev [31] concluded that overall the Muneer and Zhang is the best model among these four different correlation models for hourly diffuse radiation data at Beer Sheva, Israel. Since their model requires extra variables to be measured, we cannot test their model on our data, so we choose normalized error measures in order to make comparisons. When using the same error measures to compare the Muneer and Zhang's model with the BRL model [3], we found that the two models have similar accuracy when modelling hourly diffuse radiation. For example, the coefficient of determination of R^2 from the Muneer and Zhang's model is 0.9301 and the normalized MBE is −1.4%, whereas the BRL model the measures are 0.9628 and −3.7% respectively for hourly diffuse fraction Adelaide, Australia data. Since the evaluation is performed on separate sites, no direct comparison can be made but the results are similar in nature. Dervishi and Mahdavi [30] also assessed eight models, such as Erbs [7], Reindl [5], Orgill and Hollands [6], Lam and Li [39], Skartveit and Olseth [40], Louche et al. [41], Maxwell [26] and Vignola and McDaniels [42], for estimating diffuse fraction by using radiation data at Vienna, Austria. They found that three models, Erbs, Reindl and Orgill and Hollands performed better in obtaining estimates of diffuse fraction. Since Ridley et al. [3] showed that the BRL model performed better than the Reindl model at many locations mentioned in their paper, so the results show that the BRL model performs at least as well as one of their best performing ones. One should note in addition that the BRL model, because of its structure, is easier to implement than many of the competing models.

Instead of evaluating existing models, Li et al. [32] use a combination of different predictors, clearness index, relative sunshine duration, ambient temperature and relative humidity to estimate diffuse radiation. They compared with four other models in the literature and found that their model can perform well for estimating the monthly average daily diffuse radiation. However, they were testing their model on a longer time scale, daily, than the hourly BRL model, and so no direct comparison can be made.

From the previous studies [3,13–16], it has been shown that the multiple predictors logistic function is a suitable model for diffuse radiation. Thus, using the same method to estimate direct radiation followed naturally. This new direct radiation model was compared with the model that has been regarded as the industry standard, the Perez model [19]. It performed better overall than the Perez model. The next step was to compare this newly designed statistical direct model with the sequence of estimating the diffuse fraction with the BRL model, and subsequently calculating the direct irradiance. There proved to be no advantage in using the newly derived direct model, thus showing that there is no need to go further than using the chain of global to diffuse to direct using the BRL model.

7. Conclusion

This paper focused first on the development of models for diffuse solar radiation and then moved to discuss how to best obtain estimated hourly direct normal solar radiation. First, a logistic model for direct normal solar radiation using multiple location data was constructed. Then, the use of the logistic and Perez models in four different locations was compared. The results of four error analyses show that the logistic model performed arguably better than the Perez model. Afterwards, the BRL model was used to obtain hourly diffuse radiation and from that direct normal solar radiation. The predicted values from that chain of steps was compared with results from the logistic model. Considering all locations and error analyses, the results show that the BRL model is well equipped not to estimate the diffuse solar irradiance from the global solar, but also to go from there to estimate the direct normal irradiance.

References

[1] Lunde PJ. Solar thermal engineering. New York: John Wiley and Sons; 1979.
[2] Brunger AP. Application of an anistropic sky model to the calculation of the solar radiation absorbed by a flat plate collector. In: Proceedings Solar World Congress, Kobe. Biennial Meeting International Solar Energy Society, Kobe, Japan, September; 1989.
[3] Ridley B, Boland J, Lauret P. Modelling of diffuse solar fraction with multiple predictors. Renewable Energy 2010;35:478–83.
[4] Liu BYH, Jordan RC. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Solar Energy 1960;4(3):1–19.
[5] Reindl DT, Beckman WA, Duffie JA. Diffuse fraction correlations. Solar Energy 1990;45(1):1–7.
[6] Orgill JF, Hollands KGT. Correlation equation for hourly diffuse radiation on a horizontal surface. Solar Energy 1977;19:357.
[7] Erbs DG, Klein SA, Duffie JA. Estimation of the diffuse radiation fraction for hourly, daily and monthly average global radiation. Solar Energy 1982;4:293–302.
[8] Iqbal M. Prediction of hourly diffuse solar radiation from measured hourly global solar radiation on a horizontal surface. Solar Energy 1980;24:491–503.
[9] Bugler JM. The determination of hourly insolation on an inclined plane using a diffuse irradiance model based on hourly measured global horizontal insolation. Solar Energy 1977;19:477–91.
[10] Skartveit A, Olseth JA, Tuft ME. An hourly diffuse fraction model with correction for variability and surface Albedo. Solar Energy 1998;63(3):173–83.
[11] Gonzales JA, Calbo J. Influence of the global radiation variability on the hourly diffuse fraction correlations. Solar Energy 1999;65:119–31.
[12] Aguiar R. CLIMED Final Report. JOULE III. Project no. JOR3-CT96-0042. INETI-ITE, Dep. Renewable Energies, Lisbon; 1998. p. 53–4.
[13] Boland J, Mcarthur LC, Luther M. Modelling the diffuse fraction of global solar radiation on a horizontal surface. Environmetrics 2001;12:103–16.
[14] Boland J, Ridley B, Brown B. Models of diffuse solar radiation. Renewable Energy 2008;33(4):575–84.
[15] Lauret F, Boland J, Ridley B. Derivation of a solar diffuse fraction model in a Bayesian framework. Case Studies in Business, Industry and Government Statistics 2010;3(1):108–22.
[16] Lauret F, Boland J, Ridley B. Bayesian statistical analysis applied to solar radiation modelling. Renewable Energy 2013;49:124–7.
[17] Perez-Higueras PJ, Rodrigo P, Fernandez EF, Almonacid F, Hontoria L. A simplified method for estimating direct normal solar irradiation from global horizontal irradiation useful for CPV applications. Renewable and Sustainable Energy Reviews 2012;8(16):5529–34.
[18] Meteotest, Meteonorm 7, ⟨www.meteonorm.com⟩, 2012.
[19] Perez RR Ineichen P, Maxwell EL, Seals RD. Dynamic global-to-direct irradiance conversion models. ASHRAE Transaction, Research Series 1992:354–69.
[20] Lanini F. Division of global radiation into direct radiation and diffuse radiation. Masters Thesis, Faculty of Science, University of Bern; 2010.
[21] Duc NM. Farmers' satisfaction with aquaculture – a logistic model in Vietnam. Ecological Modelling 2008;68:525–31.
[22] Sakanoue S. Extended logistic model for growth of single-species populations. Ecological Modelling 2007;205:159–68.
[23] Banks RB. Growth and diffusion phenomena: mathematical frameworks and applications. Berlin: Spring Verlag; 1994.
[24] Jeffrey A. Advanced engineering mathematics. San Diego: Harcourt Academic Press; 2002.
[25] Thornley JHM, Shepherd JJ, France J. An open-ended logistic-based growth function: analytical solutions and the power-law logistic model. Ecological Modelling 2007;204:531–4.
[26] Maxwell EL. A quasi-physical model for converting hourly global to direct normal insolation. Solar Energy Research Institute, SERI/TR-215-3087; 1987. p. 35–46.
[27] Willmott Cj, Kenji KM. Advantages of the mean absolute error MAE over the root mean square error RMSE in assessing average model performance. Climate Research 2005;30:79–82.
[28] Massey Jr. FJ. The Kolmogorov-Smirnov test for goodness of fit. Journal of the American Statistical Association 1951;46:68–78.
[29] Espinar B, Ramirez L, Drews A, Georg Beyer H, Zarzalejo LF, Polo J, et al. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. Solar Energy 2009;83:118–25.
[30] Dervishi S, Mahdavi A. Computing diffuse fraction of global horizontal solar radiation: a model comparison. Solar Energy 2012;86(6):1796–802.
[31] Kudish AI, Evseev EG. The assessment of four different correction models applied to the diffuse radiation measured with a shadow ring using global and normal beam radiation measurements for Beer Sheva, Israel. Solar Energy 2008;82(2):144–56.
[32] Li HS, Ma WB, Wang XL, Lian YW. Estimating monthly average daily diffuse solar radiation with multiple predictors: a case study. Renewable Energy 2011;36(7):1944–8.
[33] Li HS, Bu XB, Lian YW, Zhao L, Ma WB. Further investigation of empirically derived models with multiple predictors in estimating monthly average daily diffuse solar radiation over China. Renewable Energy 2012;36:469–73.
[34] Drummond AJ. On the measurement of sky radiation. Archiv fur Meteorologie, Geophysik und Bioklimatologie Serie B 1956;7:413–36.
[35] LeBaron BA, Michalsky JJ, Perez R. A new simplified procedure for correcting shadow data for all sky conditions. Solar Energy 1990;44:249–56.
[36] Battles FJ, Alados-Arbodelas L, Olmo FJ. On shadow band correction methods for diffuse irradiance measurements. Solar Energy 1995;54:105–14.
[37] Muneer T, Zhang X. A new method for correcting shadow band diffuse solar radiation. Journal of Solar Energy Engineering 2002;128:104–17.
[38] Kudish AI, Rahima T. Unpublished results, 2005.
[39] Lam JC, Li DHW. Correlation between global solar radiation and its direct and diffuse components. Building and Environment 1996;31(6):527–35.
[40] Skartveit A, Olseth JA. A model for the diffuse fraction of hourly global radiation. Solar Energy 1987;38(4):271–4.
[41] Louche A, Notton G, Poggi P, Simonnot G. Correlation for direct normal and global horizontal irradiations on a French Mediterranean site. Solar Energy 1991;46(4):261–6.
[42] Vignola F, McDaniels DK. Correlations between diffuse and global insolation for the Pacific Northwest. Solar Energy 1984;32:161.